DOCUMENT RESUME

ED 126 381 CG 010 667

AUTHOR Jones, F. Geoffrey
TITLE The Effects of Mastery and Aptitude on Learning, Retention, and Time.
PUB DATE Apr 76
NOTE 16p.; For related document, see ED 108 981

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Academic Aptitude; Achievement; *Autoinstructional Methods; Classroom Research; Educational Research; Elementary Education; *Learning; *Retention; Speeches; *Time Factors (Learning)
IDENTIFIERS *Mastery Learning

ABSTRACT

This study used a self-instructional geography unit in 20 seventh-grade classes to assess the effects of a mastery learning procedure and aptitude upon learning, retention and time spent on studying the unit. It was found that differences between aptitude levels were increased rather than diminished when self-instructional materials were used. While self-instructional materials facilitate retention for students of high and middle aptitude, the mastery procedures did not facilitate learning and retention for low-aptitude students. Greater learning was achieved by the middle aptitude students under the mastery treatment. The lack of differences between the low aptitude mastery and non-mastery students is attributed to the lack of verbal facility that low aptitude students brought to instruction. Mastery students were found to use considerably more time in learning the material than did non-mastery students. The implications of these findings are discussed in relation to previous research and practical application. (SJL)

The Effects of Mastery and Aptitude on Learning, Retention and Time.

F. Geoffrey Jones

Memorial University of Newfoundland

Paper presented at the A.E.R.A. Conference, San Francisco, April 19-23, 1976 in the symposium titled "Is a Mastery Learning - C.B.E. Model Appropriate to the Social Studies?" - A critique of four studies.

Mastery learning is a teaching-learning procedure that claims to improve learning achievement for up to ninety-five percent of a given student population (Bloom, 1968). To facilitate such achievement Bloom proposed procedures "whereby each student's instruction and learning can be managed within the context of ordinary group-based classroom instruction, as to promote his fullest development." By manipulating the amount of time spent on learning, all students, and particularly low-performing students, could spend differential time allotments to promote maximum learning. However, the implication here that mastery learning closes the learning gap between students of varying aptitude requires closer scrutiny.

The Bloom hypothesis that mastery learning can overcome aptitude differences is contrary to the mass of psychological evidence which indicates that most treatments are insufficient to overcome differences in aptitude (De Cecco, 1968), and that methods of teaching share the common result of ineffectiveness (Wallen and Travers, 1963). Furthermore, Bloom mastery procedures are class-based rather than individual-paced. In the Bloom procedure, the learning progress of the high aptitude student is retarded by the withholding of additional learning tasks. Instead, the high achiever serves as tutor or teacher aide in assisting the lower-performing or slower student.

By contrast, individual-paced instruction, whether of the earlier Winnetka-type (Washburne, 1922) or the more recent I.P.I.-type (Glaser, 1968), has shown superior performance by students of high aptitude when compared to middle and low aptitude students when sequential and hierarchically-structured learning materials have been used. However, how might achievement and aptitude be affected by a mastery learning procedure that utilizes a self-instructional format with content whose organization is often left to the arbitrary whim of the classroom teacher?

In a critical analysis of the state of the art and quality of research, Mitchell (1975) concluded that much mastery research is based upon crude comparisons of a mastery group with a non-mastery(group with a non-mastery group)often with ex post facto comparisons. Thus, while mastery learning procedures have generally been reported as superior to non-mastery procedures (Kim, 1969, 1970; Block, 1970; Lee, 1971), studies controlling the time variable have shown increasing efficiency over a series of sequenced learning work (Merrill, Barton, and Wood, 1970; Block, 1970), and studies by Block (1970), Kersh (1970), Romberg, Shepler, and King (1970) and Wentling (1970) found increased retention where Bloom class-biased procedures were used. The remaining studies using individualized procedures (Green, 1969; Gentile, 1970) indicate superior performance but are insufficient to answer problems dealing with aptitude, time, or retention. Further, only two studies were located that used content from the social science disciplines (Gaines, 1971; Tierney, 1973).

Consequently a study was designed that used a self-instructional geography unit Functions of Cities (Jones, 1974) that would assess the effects of a mastery learning procedure and aptitude upon learning, retention, and time spent on studying the unit. The generic research question of this study was:

> If a self-instructional mastery learning procedure is used in teaching a geography unit at the grade seven level, will the average achievement of students at three levels of aptitude be significantly different?

Method

Two treatment procedures were employed in this study. They were a mastery (T_1) and a non-mastery (T_2) learning procedure. As the content of the materials Functions of Cities (Jones, 1974) was the same for both treatments, the focus of the study was on the manipulation of various components within the treatment structures. Table 1 contains the outline used in conceptualizing each of the components.

TABLE 1

Layout of Mastery and Non-Mastery Procedures

	Treatment 1 (Mastery)	Treatment 2 (Non-Mastery)
Presentation		
Narrative	X	X
Student Workbook Activities	X	X
Diagnosis		
Review Test One	X	X
Correction	X	0
Feedback	X	0
Remediation		
Prescriptive Review	X	0
Specific Practice	X	0
General Review	X	0
Diagnosis		
Review Test Two	X	0
Correction	X	0
Summative Test		
(Administered to all students at the conclusion of the unit)	X	X
Weekly Class Discussion	X	X

The X's indicate the components that were used in the procedures while the 0's indicate those components not used.

Instrumentation

Two instrument types were constructed. They were the Review Tests and the Summative Test (Geography Achievement Test).

Review Tests: Each chapter of the workbook Functions of Cities contained a review test. A review test measured the amount learned in each chapter. Each review test contained twenty items. Items were written in three forms: (1) three-foil multiple choice; (2) true or false; and (3) completion. Items tested recall, application, and transfer of cognitive knowledge.

Review tests used in both treatments were exactly the same. The non-mastery (T_2) treatment contained one review test at the conclusion of each chapter. The mastery (T_1) treatment contained two review tests. The second review test for the mastery treatment contained exactly the same items; however they were placed in a different order. Students in the mastery treatment were required to reach the 75% criterion level before progressing to the next unit of work.

Summative Test: The summative test was in two parts. The first part was a 40 item, four-option, multiple-choice test, while the second part was a 24-item retrieval chart completion test. The total 64-item test was designed to measure the students' knowledge of facts, concepts, and generalizations presented in the treatment unit.

Using the Analysis of Item and Test Homogeneity (A.N.L.I.T.H.) computer program the 40-item multiple choice test obtained a reliability coefficient of .89 and the 24-item completion test obtained a reliability coefficient of .95.

Subjects

Twenty grade seven classes (N=539) were obtained from the Savannah-Chatham County Public Schools in Savannah, Georgia.

Procedure

A 3x10x2 aptitude by classes-nested-within-treatments, by treatments, multivariate analysis of variance (M.A.N.O.V.A.) using three measures of effect was employed with the posttest data of this study . This experimental design was used to determine if the differences between the mastery and non-mastery treatments produced differences ($P<.05$) in the average effects which were not the same at the high, middle and low aptitude levels.

The B.M.D. 12V program (Biomedical Computer Programs, 1973) was used for the above analysis. This program was used because it can perform multivariate and univariate analyses of variance for any hierarchical design with cells that contain equal N's, including nested, partially nested and partially crossed, and fully crossed designs.

If the multivariate analysis was found to be statistically significant an a priori decision was made to follow up by testing the univariate interaction hypotheses and no main effects would be tested for. However, if there was no statistical significance on the multivariate interaction hypothesis, each of the multivariate main effects was to be tested. If these were statistically significant then each effects measure would be measured at the univariate level. Duncan's Multiple Range Test was the appropriate post hoc test for statistically-significant outcomes for the univariate analyses (Edwards, 1968), while the Bonferroni t statistic was the appropriate post hoc test for simple effects (Marascuilo and Levin, 1970).

Results

The multivariate interaction hypothesis was not significant. However, the multivariate main effects hypotheses (treatment and aptitude) were significant (.001). Univariate analyses were subsequently conducted for the three measures of effect (learning, retention, and time). The results have been tabulated in Tables 2 and 3.

Table 2

Summary of Multivariate and Univariate Tests of Significance: Interaction and Main Effects

Statistical Hypotheses (Null)	F	Level of Significance
There are no differences:		
I. Between vectors (MANOVA) of learning, retention, and times-to-testing;		
1. Interaction: treatment by aptitude	1.02	N.S.
2. Main Effects: treatment	14.82	.001
3. Main Effects: aptitude	14.99	.001
II. Learning (ANOVA): mean differences for interaction and main effects;		
4. Interaction: treatment by aptitude	2.87	N.S.
5. Main Effects: treatment	2.99	N.S.
6. Main Effects: aptitude	56.39	.05
III. Retention (ANOVA): mean difference for interaction and main effects;		
7. Interaction: treatment by aptitude	2.07	N.S.
8. Main Effects: treatment	16.28	.05
9. Main Effects: aptitude	66.74	.05
IV. Times-to-Testing (ANOVA): mean differences for interaction and main effects;		
10. Interaction: treatment by Aptitude	0.34	N.S.
11. Main Effects: treatment	26.60	.05
12. Main Effects: aptitude	0.19	N.S.

Table 3

Summary of Tests of Significance for Simple Effects: Comparisons of Aptitude Levels Across Treatments

Statistical (Null) Hypotheses	Mean Score Mastery Treatment	Mean Score Non-Mastery Treatment	Mean Difference	Level of Significance
There are no differences:				
I. Learning: treatment means across aptitude levels.				
13. High	38.28	34.40	3.87	N.S.
14. Middle	30.26	23.83	6.43	.05
15. Low	15.22	17.56	-2.34	N.S.
II. Retention: treatment means across aptitude levels.				
16. High	40.82	33.25	9.56	.05
17. Middle	29.67	20.82	8.85	.05
18. Low	16.67	14.86	1.81	N.S.
III. Times-to-Testing: treatment means across aptitude levels.				
(19) High	537.20*	447.67*	59.53*	.05
(20) Middle	556.46*	447.55*	78.91*	.05
(21) Low	561.33*	472.96*	88.38*	.05

*Expressed in minutes

A statistically-significant difference was found between the mean scores for the three levels of aptitude on the learning measure (posttest). The Duncan Multiple Range Test reported the following outcomes.

Table 4

Learning: Summary of Results of the Duncan Multiple Range Test at the .05 Level of Significance for Aptitude Effect.

Pairwise Comparisons	Significance
1 vs 2	.05
1 vs 3	.05
2 vs 3	.05

The high aptitude group learned more than the middle and low aptitude groups, and the middle aptitude group learned more than the low aptitude group.

The mean scores for treatments on the retention measure were found to be statistically significant in favour of the mastery treatment. As well, statistical significance was found between the mean scores for the three levels of aptitude on the retention measure (delayed posttest). The Duncan Multiple Range Test reported the following outcomes.

Table 5

Retention: Summary of Results of the Duncan Multiple Range Test at the .05 Level of Significance for Aptitude Effect.

Pairwise Comparisons	Significance
1 vs 2	.05
1 vs 3	.05
2 vs 3	.05

The high-aptitude group retained more than the middle-and low-aptitude groups and the middle-aptitude group retained more than the low-aptitude group.

The mean scores for treatments on the time measure were statistically significant in favour of the non-mastery treatment. The non-mastery treatment group took significantly less time than the mastery group.

A Bonferroni t statistic test was subsequently used to determine if statistically-significant differences occurred at each level of aptitude between treatments for learning, retention and time. At .10 level of significance and with 36 degrees of freedom at difference as large as 5.89 was needed to indicate a statistically-significant difference.

A statistically-significant difference for learning was produced between the middle aptitude mastery and non-mastery group in favour of the mastery group while the high and middle aptitude mastery groups produced statistically-significant differences for the amount of learning retained. Statistical significance was obtained in favour of the non-mastery groups at the three levels of aptitude for the amount of time used to completion of the study. However, a difference as large as 17.80 with 36 degrees of freedom was required for a significant difference.

Discussion

This study found that differences between aptitude levels were increased rather than diminished when self-instructional materials were used. High aptitude students learned and retained more of the geography unit than middle or low aptitude students, while middle aptitude students learned and retained more of the geography unit than low aptitude students. These results suggest that achievement was a function of the capacities and talents for learning that students of varying aptitude brought to the instruction.

The analysis of simple effects of treatments across each level of aptitude found that the mastery treatment facilitated greater retention for the high and middle aptitude students, and greater learning for the middle aptitude students. This was accomplished due to the feedback correction procedures required of the mastery students and the increased time that these procedures required of the mastery students for relearning. This result is consistent with that of Fishburne (1971) who used a programmed and non-programmed text. He found that exposure to the programmed text increased learning and retention but took more time across levels of reading. He attributed increased student learning to the extra time taken with the materials. Therefore, it would appear that self-instructional materials at least facilitate retention for students of high and middle aptitude students. However, the mastery procedures did not facilitate learning and retention for low-aptitude students.

Low aptitude mastery students neither learned nor retained the geography material more than low aptitude non-mastery students. The low aptitude students used in this study obtained very low reading scores as measured by the Iowa Test of Basic Skills. When converted to grade equivalent scores the low aptitude mastery and non-mastery students were reading at approximately fourth grade level. This is almost four grade levels below actual classroom level and at least two grades below the Grade 6 reading level of the materials Functions of Cities used in the study. Therefore, the lack of differences between the low aptitude mastery and non-mastery students can be explained by the lack of verbal facility that low aptitude students brought to instruction. This was particularly evident in the scores obtained on the 40-item multiple-choice and the 24-item recall chart. The low aptitude mastery and non-mastery students consistently scored lower than the middle and high aptitude groups on the 40-item multiple choice test and often did not start the 24-item recall test. This strongly suggests that the strength of learning by low aptitude students was indeed low. Another factor that reinforces this position is that there was only a one-chapter difference between high and low aptitude students at the completion of

instruction. This suggests that low aptitude students did not spend the necessary time in relearning the material necessary to improve their learning. The difficulty of the material due to their inherent reading and vocabulary deficiencies probably caused frustration in learning and reduced their task orientation. Therefore, the materials Functions of Cities were probably too difficult for low aptitude students.

The review of the nine studies comparing mastery to non-mastery strategies revealed that two were below the college level, three used self-instructional materials, and none used social science materials. Within this context, all studies reported that mastery facilitated learning more than a non-mastery treatment. The emphasis of research was at the university or college level where students used could not be considered a representative sample of normal classroom conditions.

The results of the present study indicate that when self-instructional mastery procedures are used they do not facilitate greater posttest average performance than non-mastery procedures. The findings are contrary to Moore, Mahan and Ritts (1968), Green (1969), and Gentile (1970). These researchers used self-paced procedures. However, they used content that is sequential by nature (math and science content) and each learning task was contiguous with the next. This study used geography materials organized in a specific sequence devised by the researcher. However, the materials were constructed and organized around two major generalizations and this scheme was followed through each of the chapters. The results of the present study apply to the materials and students in this study but it is reasonable to suppose that similar results would be obtained if the same materials were used with students who contained similar contextual characteristics.

The literature concerning retention (Block, 1970; Kersh, 1970; Romberg, Shepler, and King, 1970; and Wentling, 1970) found that retention is facilitated when group-paced instruction is used with correction and feedback. This study found that when self-instructional geography materials were used mastery procedures

facilitated greater retention than non-mastery procedures as measured by the delayed posttest. Therefore, this would suggest that the correction feedback procedures, either group-paced or self-instruction, facilitated greater retention of original learning.

The literature review showed that only two studies reported the time variable (Merrill, Barton, and Wood, 1970; Block, 1970). Both studies indicated that learning became increasingly efficient over a series of sequenced learning units in class-paced instruction. This study did not support these findings. Mastery students used considerably more time to learn the material than non-mastery students. These time differentials also increased when comparisons were made between aptitude levels. Therefore, the results of this study would suggest that self-paced mastery instruction requires more time than self-paced, non-mastery instruction or class-paced instruction.

BIBLIOGRAPHY

Block, J. H. The effects of various levels of performance on selected cognitive, affective and time variables. Unpublished Ph.D. dissertation, University of Chicago, 1970.

Bloom, B. S. Learning for mastery, Evaluation Comment 1, No. 2 (1968).

DeCecco, J. P. The psychology of learning and instruction: Educational psychology. Englewood Cliffs, New Jersey: Prentice Hall, 1968.

Dixon, W. J. (Ed.), Biomedical Computer Programs. Los Angeles: University of California Press, 1973.

Fishburne, R. P.,Jr. The comparison of a programmed and a non-programmed text on evolution for the fifth grade. Unpublished Master's thesis, Georgia Southern College, 1971.

Gaines, W. G. An application of John B. Carroll's model of school learning to the teaching of anthropology. Unpublished doctoral dissertation, University of Georgia, 1971.

Gentile, J. R. A mastery strategy for introductory educational psychology. Unpublished material, State University of New York at Buffalo, Department of Educational Psychology, 1970.

Glaser, R. Adapting the elementary curriculum to individual performance. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton: Educational Testing Service, 1968, 3-36.

Green, B. A. A self-paced course in freshman physics. Cambridge, Massachusetts: Massachusetts Institute of Technology, Education Research Center, 1969.

Jones, F. G. Functions of cities. Geography Curriculum Project, Publication No. 74-1, University of Georgia.

Kersh, M. E. A strategy for mastery learning in fifth-grade arithmetic. Unpublished Ph. D. dissertation, University of Chicago, 1970.

Kim, H. Learning rates, aptitudes, and achievements. Unpublished Ph.D. dissertation, University of Chicago, 1970.

Kim, H. A study of the Bloom strategies for mastery learning. Seoul: Korean Institute for Research in the Behavioral Sciences, 1970.

Lee, Y. P. Interaction improvement studies on the mastery learning project - final report on mastery learning program, April - November, 1971. Research Center, Seoul National University November, 1971.

Lindquist, E.F., & Hieronymus, A.N. Iowa tests of basic skills, Forms 5 and 6. Boston, Mass.: Houghton Mifflin Co., 1971.

Maracuilo, L.A., & Levin, J.R. Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: The elimination of type IV errors. American Educational Research Journal, 1970, 7(3), 397-421.

Merrill, M.D., Barton, K., & Wood, L.E. Specific review in learning a hierarchical imaginary science, Journal of Educational Psychology, 61, 1970, 102-9.

Mitchell, J.J. A critical analysis of the research on mastery learning. Unpublished Master's thesis, University of Georgia, 1974.

Moore, J.W., Mahan, J.M., & Ritts, C.A. An evaluation of the continuous progress concept of instruction with university students. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, 1968.

Romberg, A. Shepler, J., & King, I. Mastery Learning and Retention. Technical Report No. 151, Wisconsin Research and Development Center for Cognitive Learning. The University of Wisconsin, Madison, Wisconson, 1970.

Tierney, M.L. A Comparison of Feedback/Correction Procedures used in Mastery Learning Strategies. Unpublished Master's thesis, University of California, Santa Barbara, 1973.

Wallen, N.E., & Travers, R.M.W. "Analysis and Investigation of Teaching Methods", Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Washburne, C.W. Educational Measurement as a Key to Individualizing Instruction and Promotions. Journal of Educational Research, 5 (1922), 195-206.

Wentling, T.L. Mastery Versus non-mastery Instruction Varying Test Item Feedback Instruments. Journal of Educational Psychology (in press), 1972.